A SOLUTION TO A COUNTABLE SYSTEM OF EQUATIONS ARISING IN MARKOVIAN DECISION PROCESSES

by

Cyrus Derman and Arthur F. Veinott, Jr.

TECHNICAL REPORT NO. 89

July 7, 1966

Supported by the Army, Navy, Air Force, and NASA under
Contract Nonr-225(53) (NR-042-002)
with the Office of Naval Research*

Gerald J. Lieberman, Project Director

*This research was partially supported by the Office of Naval Research under Contract Nonr-225(77) (NR-347-010).

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Nontechnical Summary

Let $X_o, X_1, \ldots$ be a sequence of non-negative integer valued random variables with the property that

$$\Pr(X_{n+1} = j \mid X_o = x_o, \ldots, X_{n-1} = x_{n-1}, X_n = i) = p_{ij}$$

for all $i, j, x_o, \ldots, x_n, n$. The collection of random variables $\{X_n\}$ is called a Markov chain and the p_{ij} are called transition probabilities. We refer to X_n as the state of the process at time n. Let w_i be the cost incurred at time n if the process is in state i at that time. Consider the system of equations

$$(1) \qquad g + v_i = w_i + \sum_{j=o}^{\infty} p_{ij} v_j, \quad i = 0, 1, \ldots$$

in the unknown variables $g, v_o, v_1, \ldots$. Such a system arises in connection with constructing optimal rules for controlling Markovian decision processes. Also the numbers $g, v_o, v_1, \ldots$ are of interest in their own right. Often g is the long run expected average cost and $v_i - v_j$ is the limit, as $n \to \infty$, of the difference between expected total cost during times $0, 1, \ldots, n$ given that the process starts in states i and j respectively.

We show in this paper that one solution to the system (1) is given by

$$(2) \qquad g = \frac{c_{oo}}{m_{oo}} \quad \text{and} \quad v_i = c_{io} - g m_{io}, \quad i = 0, 1, \ldots$$

provided that the expected time m_{io} required to go from state i to state 0 is finite and that the expected cost c_{io} incurred during that time is also finite, $i = 0, 1, \ldots$. Notice that $v_o = 0$.

As an illustration of the above ideas, consider a single item inventory model in which the demands in periods $1, 2, \ldots$ are independent. A demand of size one occurs with probability p, $0 < p < 1$, and a demand of size zero occurs with probability $1 - p$. Let X_n denote the stock on hand at the beginning of period n. An order for one unit is placed in period n with immediate delivery if $X_n = 0$; otherwise, no order is placed in period n. There is a unit cost h for each unit of stock on hand after ordering in a period. There is a cost K for placing an order in a period. Under these assumptions the nonzero transition probabilities are $p_{oo} = p$, $p_{o1} = 1 - p$, $p_{ii} = 1 - p$, and $p_{i,i-1} = p$, $i = 1, 2, \ldots$. Also $w_o = K + h$ and $w_i = hi$, $i = 1,2,\ldots$. Thus the system (1) becomes

$$g + v_o = K + h + pv_o + (1 - p)v_1$$

$$g + v_i = ih + pv_{i-1} + (1 - p)v_i, \; i = 1, 2, \ldots .$$

The solution given in (2) is

$$g = pK + h ,$$

$$v_i = \frac{hi(i-1)}{2p} - Ki, \; i = 0, 1, \ldots .$$

Thus g is here the long run expected average cost under the indicated ordering policy. Also v_i is the limit, as $n \to \infty$, of the amount by which the expected cost in periods $0, 1, \ldots, n$ starting with i units of stock on hand exceeds that starting with no stock on hand.

A SOLUTION TO A COUNTABLE SYSTEM OF EQUATIONS

ARISING IN MARKOVIAN DECISION PROCESSES

by

Cyrus Derman
Columbia University and Stanford University

and

Arthur F. Veinott, Jr.
Stanford University

Let $\{X_n\}$, $n = 0, 1, \ldots$, be a Markov chain having a state space consisting of the non-negative integers and having stationary transition probabilities $\{p_{ij}\}$. Let $\{w_i\}$, $i = 0, 1, \ldots$, be a sequence of real numbers. Consider the system of equations

$$(1) \qquad g + v_i = w_i + \sum_{j=0}^{\infty} p_{ij} v_j, \quad i = 0, 1, \ldots ,$$

in the unknown variables $\{g, v_0, v_1, \ldots\}$. In [2], the system (1) arises in connection with conditions for the existence and construction of optimal rules for controlling a Markovian decision process. For a finite state space existence of solutions to (1) is guaranteed by the condition that the Markov chain have at most one ergodic class of states. (See [3].) In this note we give conditions ensuring the existence (Theorem 1) and uniqueness (Theorem 2) of solutions to (1).

Let

$$Z_n(j) = \begin{cases} 1, & \text{if } X_n = j \text{ and if } X_m \neq 0 \text{ for } 0 < m \leq n \\ 0, & \text{otherwise} \end{cases}$$

$$j, n = 0, 1, \ldots ,$$

$${}_0p^*_{ij} = E\left(\sum_{n=0}^{\infty} Z_n(j) \,\middle|\, X_0 = i\right), \quad i, j = 0, 1, \ldots ,$$

and

$$m_{io} = \sum_{j=o}^{\infty} {}_{o}p^{*}_{ij}, \quad i = 0, 1, \dots .$$

If the last series converges absolutely, then m_{io} is the mean first passage time from i to 0 and we say m_{io} is finite. If the m_{io} are all finite, as we assume throughout, then state 0 is positive recurrent and there is only one recurrent class.

$$\text{Let} \quad Y_n = \sum_{j=o}^{\infty} w_j Z_n(j) \quad \text{and} \quad c_{io} = E\left(\sum_{n=o}^{\infty} Y_n \middle| X_o = i\right) .$$

By an obvious generalization of Theorem 5 in [1, p. 81] we get $c_{io} = \sum_{j=o}^{\infty} {}_{o}p^{*}_{ij} w_j$ provided the series is absolutely convergent. If the series is absolutely convergent we say c_{io} is finite. In applications w_i is often the cost incurred when in state i so c_{io} is then the expected cost during a first passage from i to o.

Theorem 1 (Existence)

If the numbers m_{io} and c_{io}, $i = 0, 1, \dots$, are finite, then the numbers

$$(2) \qquad g = \frac{c_{oo}}{m_{oo}} \quad \text{and} \quad v_i = c_{io} - gm_{io}, \quad i = 0, 1, \dots$$

satisfy (1) and $\sum_{j=o}^{\infty} p_{ij} v_j$ converges absolutely, $i = 0, 1, \dots$.

Proof:

Let $w^{*}_i = w_i - g$ and $Y^{*}_n = \sum_{j=o}^{\infty} w^{*}_j Z_n(j)$. Then for $i = 0, 1, \dots$

$$v_i = E\left(\sum_{n=o}^{\infty} Y_n^* \middle| X_o = i\right)$$

$$= w_i^* + \sum_{n=1}^{\infty} \sum_{j=o}^{\infty} E(Y_n^* | X_o = i,\ X_1 = j) p_{ij}$$

$$= w_i^* + \sum_{j=o}^{\infty} \sum_{n=1}^{\infty} E(Y_n^* | X_o = i,\ X_1 = j) p_{ij}$$

$$= w_i^* + \sum_{j=o}^{\infty} p_{ij} v_j$$

so (1) holds. The interchange of expectation and summation is justified since the finiteness of the m_{io} and c_{io} imply that $\sum_{n=o}^{\infty} E(|Y_n^*| \, | X_o = i) < \infty$. This in turn implies that the series above are absolutely convergent so the interchange of summations is also justified.

<u>Theorem 2</u> (Uniqueness)

If the numbers m_{io} and c_{io}, $i = 0, 1, \dots$, are finite, if $\sum_{j=o}^{\infty} {}_oP_{ij}^* \left(c_{jo} - \frac{c_{oo}}{m_{oo}} m_{jo}\right)$, $i = 0, 1, \dots$ converges absolutely, and if $\{g, v_o, v_1, \dots\}$ is a sequence with $\sum_{j=o}^{\infty} {}_oP_{ij}^* v_j$, $i = 0, 1, \dots$, converging absolutely, then $\{g, v_o, v_1, \dots\}$ satisfies (1) if and only if there is a real number r such that

(3) $$g = \frac{c_{oo}}{m_{oo}} \quad \text{and} \quad v_i = c_{io} - g m_{io} + r,\ i = 0, 1, \dots .$$

<u>Proof</u>:

It is immediate from the hypotheses and Theorem 1 that $\{g, v_o, v_1, \dots\}$ defined in (3) satisfies (1) and $\sum_{j=o}^{\infty} {}_oP_{ij}^* v_j$ converges absolutely as well as $\sum_{j=o}^{\infty} p_{ij} v_j$. Let $\{g', v_o', v_1', \dots\}$ be

any other solution to (1) with $\sum_{j=0}^{\infty} {}_0p^*_{ij}v'_j$ converging absolutely for $i = 0, 1, \ldots$. Hence $\sum_{k=0}^{\infty} p_{ik}v'_k$ is absolutely convergent. Now pre-multiplying both sides of (1) by $\pi_i \equiv \frac{{}_0p^*_{oi}}{m_{oo}}$, summing over $i = 0, 1, \ldots$, using the relations $\sum_{i=0}^{\infty} \pi_i = 1$ and $\pi_j = \sum_{k=0}^{\infty} p_{kj}\pi_k$, $j = 0, 1, \ldots$, and the fact that the interchange of summations is justified, we get $g' = \sum_{i=0}^{\infty} \pi_i w_i$ which is independent of $\{v'_o, v'_1, \ldots\}$. Thus since $\{g, v_o, v_1, \ldots\}$ satisfies (1) we must have $g = g'$.

Letting $\Delta_i = v'_i - v_i$, $i = 0, 1, \ldots$, we get from (1) on subtracting one system from the other that

$$\Delta_i = \sum_{j=0}^{\infty} p_{ij}\Delta_j, \quad i = 0, 1, \ldots . \tag{4}$$

Let $p^n_{ij} = Pr(X_n = j | X_o = i)$. Evidently for $N = 1, 2, \ldots$,

$$\sum_{n=1}^{N} p^n_{ij} \leq {}_0p^*_{ij} + (N-1)\,{}_0p^*_{oj}, \quad j = 0, 1, \ldots$$

so

$$\frac{1}{N}\sum_{n=1}^{N} p^n_{ij}\,|\Delta_j| \leq [{}_0p^*_{ij} + {}_0p^*_{oj}]|\Delta_j|, \qquad j = 0, 1, \ldots . \tag{5}$$

Since the series on the right side of (5) converges absolutely by hypothesis, and $\lim_{N \to \infty} \frac{1}{N}\sum_{n=1}^{\infty} p^n_{ij} = \pi_j$, we get from the dominated convergence theorem that

(6) $$\lim_{N\to\infty} \sum_{j=o}^{\infty} \frac{1}{N} \sum_{n=1}^{\infty} p_{ij}^{n} \Delta_j = \sum_{j=o}^{\infty} \pi_j \Delta_j \ .$$

Since from (5), $\sum_{j=o}^{\infty} p_{ij}^{n} \Delta_j$ converges absolutely we can iterate (4), yielding

(7) $$\Delta_i = \sum_{j=o}^{\infty} p_{ij}^{n} \Delta_j, \quad i = 0, 1, \ldots ; \quad n = 1, 2, \ldots .$$

Hence on substituting (7) into (6)

$$\Delta_i = \sum_{j=o}^{\infty} \pi_j \Delta_j, \quad i = 0, 1, \ldots .$$

Thus Δ_i is independent of i, which completes the proof.

Example:

If the sequences $\{m_{io}\}$ and $\{w_i\}$, $i = 0, 1, \ldots$, are bounded, then so is the sequence $\{c_{io}\}$, $i = 0, 1, \ldots$, since $|c_{io}| \leq \sup_{k,j} m_{ko} |w_j|$. Thus Theorem 1 applies and in addition the solution to (1) given in (2) is bounded. This result is used in [2].

We remark that since

$$\sum_{j=o}^{\infty} {}_{o}p_{oj}^{*} |u| \geq {}_{o}f_{ok} \sum_{j=o}^{\infty} {}_{o}p_{kj}^{*} |u_j|$$

where

$${}_{o}f_{ok} = \Pr\left(\sum_{n=o}^{\infty} Z_n(k) > 0 \,|\, X_o = 0 \right) > 0 \ ,$$

$\sum_{j=0}^{\infty} {}_0p^*_{kj}|u_j|$ is absolutely convergent for every recurrent state k provided that $\sum_{j=0}^{\infty} {}_0p^*_{0j}|u_j|$ is absolutely convergent. Thus the hypotheses of Theorems 1 and 2 could have been stated only for state 0 and the transient states.

References

[1] Chung, K. L. (1960), Markov Chains with Stationary Transition Probabilities, Springer, Berlin.

[2] Derman, Cyrus (1966), "Denumerable State Markovian Decision Processes - Average Cost Criterion," (To Appear in Ann. Math. Stat.).

[3] Howard, Ronald (1960), Dynamic Programming and Markov Processes, John Wiley, New York.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Stanford University Department of Statistics Stanford, California	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP
3. REPORT TITLE A Solution to a Countable System of Equations Arising in Markovian Decision Processes.	
4. DESCRIPTIVE NOTES (Type of report and inclusive dates) Technical Report	
5. AUTHOR(S) (Last name, first name, initial) Derman, Cyrus Veinott, Arthur F., Jr.	

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
July 7, 1966	10	3

8a. CONTRACT OR GRANT NO. Contract Nonr-225(53) b. PROJECT NO. NR-042-002 c d	9a. ORIGINATOR'S REPORT NUMBER(S) Technical Report No. 89 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) None
10. AVAILABILITY/LIMITATION NOTICES Distribution of this document is unlimited.	
11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Logistics and Mathematical Statistics Branch Office of Naval Research Washington, D. C. 20360

13. ABSTRACT

A countable system of equations arising in Markovian decision processes is studied. Conditions are given ensuring the existence and uniqueness of an explicit solution to the system.

DD FORM 1473 JAN 64

UNCLASSIFIED

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Markov chains Dynamic Programming Markov decision processes						

INSTRUCTIONS

1. ORIGINATING ACTIVITY: Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2*a*. REPORT SECURITY CLASSIFICATION: Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2*b*. GROUP: Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. REPORT TITLE: Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. DESCRIPTIVE NOTES: If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. AUTHOR(S): Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. REPORT DATE: Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7*a*. TOTAL NUMBER OF PAGES: The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7*b*. NUMBER OF REFERENCES: Enter the total number of references cited in the report.

8*a*. CONTRACT OR GRANT NUMBER: If appropriate, enter the applicable number of the contract or grant under which the report was written.

8*b*, 8*c*, & 8*d*. PROJECT NUMBER: Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9*a*. ORIGINATOR'S REPORT NUMBER(S): Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9*b*. OTHER REPORT NUMBER(S): If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. AVAILABILITY/LIMITATION NOTICES: Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through
_________________________________."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through
_________________________________."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through
_________________________________."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. SUPPLEMENTARY NOTES: Use for additional explanatory notes.

12. SPONSORING MILITARY ACTIVITY: Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. ABSTRACT: Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS)*, *(S)*, *(C)*, or *(U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. KEY WORDS: Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.